

MCDoseE 2.0 A new Markov Chain Monte Carlo program for ESR dose response curve fitting and dose evaluation



Renaud Joannes-Boyau^{a,*}, Mathieu Duval^b, Thomas Bodin^c

^a Southern Cross GeoScience, Southern Cross University, Lismore, NSW, 2480, Australia

^b Australian Research Centre for Human Evolution (ARCHE), Environmental Futures Research Institute, Griffith University, Nathan, QLD, 4111, Australia

^c Laboratoire Géologie de Lyon: Terre, Planètes, Environnement, Université de Lyon and Ecole Normale Supérieure de Lyon, UMR CNRS 5276, F-69622, Villeurbanne, France

A B S T R A C T

This study presents MCDoseE 2.0, a new fitting program for ESR dating dose response curve (DRC) fitting and dose calculation. The standalone software was specifically designed to remove assumed data weighting, and instead to obtain a full probabilistic solution of the DRC by propagating the uncertainties associated with the measured ESR intensities. It uses a non-linear Bayesian framework, specifically a Markov Chain Monte Carlo (MCMC) scheme based on the Metropolis-Hastings algorithm, where the solution is a probability distribution for the equivalent dose, according to the precision of the measurements.

In this paper, we investigate the capabilities and limitations of MCDoseE 2.0 by comparing our results to those obtained with OriginPro 9.1[®], a proven and commonly used commercial software package. The two programs were evaluated against both known-dose samples and random archaeological tooth enamel and quartz samples, using three commonly used DRC fitting functions. We found that both programs provide highly consistent results. When comparing the dose estimates obtained by both programs we found that 90% of the solutions are statistically indistinguishable regardless of the data weighting assumption used in OriginPro. We also found that MCDoseE 2.0 offers an increased precision on the ending results compared to the commercial software, as long as each measured ESR uncertainty remains within 2-sigma range of the mean error value of all measured ESR uncertainties of the dataset. The accuracy of the fitting results given by MCDoseE 2.0 are undeniably dependent on the measurement accuracy, and emphasises the need of a proper assessment of the experimental errors in the ESR intensities.

A copy of the program is available in Supplementary information, and some basic instructions for its use are provided, as well as recommendations to ensure reliable and accurate fitting results.

1. Introduction

In ESR dating/dosimetry, the fitting of the experimental data points is a key step for obtaining accurate values for the equivalent dose (D_E). Since the first application of the method in the early 1970s, the use of an increasing number of experimental points at higher irradiation doses have made the ESR dose response curves (DRCs) more complex, invalidating a linear fitting approach. Consequently, several fitting algorithms have been tested over recent decades to model saturating exponential dose response curves (e.g. Hayes et al., 1998; Grün and Brumby, 1994 and references therein). This is usually done either by researchers developing their own program or by using a commercially available software package. While it seems preferable to choose the first option, as it offers the possibility to design a program that perfectly

addresses the research purpose, the advanced knowledge required in several fields such as computer programming, mathematics and statistics renders the task complicated and time consuming. Hence, the predilection for commercial software use in ESR dating (e.g. OriginPro, Kaleidagraph). Nonetheless, even with a detailed user manual, most commercial software perform frequently as “black boxes”, with limited understanding of the fitting process or control over the results.

In a recent paper, Duval and Grün (2016) compared the performances of two distinct fitting programs using the Single Saturating Exponential (SSE) function for D_E reconstruction of fossil enamel. The first was a non-commercial program, FIT-RSES, had been specifically designed by R. Grün for the above-mentioned purpose and has been continuously improved over the last two decades (e.g. Grün and Brumby, 1994; Grün, 2000, 2002, 2006). Succinctly, in FIT-RSES the

* Corresponding author.

E-mail address: renaud.joannes-boyau@scu.edu.au (R. Joannes-Boyau).

Table 1
Equations of the various fitting functions used in the present work.

Function	Equation	Fitted param.
SSE	$I(D) = I_{sat} * (1 - e^{-(D+D_E)/D_0})$	(3): I_{sat} , D_0 , D_E .
DSE	$I(D) = I_1 * (e^{-(D+D_E)/D_1}) + I_2 * (e^{-(D+D_E)/D_2})$	(5): I_1 , I_2 , D_1 , D_2 , D_E .
EXP + LIN	$I(D) = I_1 * (1 - e^{-(D+D_E)/D_0}) + m * (D + D_E)$	(4): I_1 , m , D_0 , D_E .

ESR DRC is fitted by linearization of the SSE function, which offers the advantage of a simple and fast optimization of the parameters (see further details in Apers et al., 1981; Grün and Brumby, 1994). The second program benchmarked in their study was the commercial software developed by OriginLab Corporation, which is widely used for data analysis (e.g. peak analysis, curve fitting, statistics, signal processing ...) in several disciplines. This software allows non-linear fittings based on the Levenberg-Marquardt algorithm. The graphical interface makes it easy to use, which has led Origin to be the preferred DRC fitting tool in the ESR dating community (e.g. Duval et al., 2009; Han et al., 2012; Hoffmann et al., 2003; Kinoshita et al., 2008; Küçükuysal et al., 2011). Duval and Grün (2016) demonstrated that the two fitting methods provided highly consistent results with the SSE function without introducing any identifiable bias. For equivalent dose (D_E) values ranging around 120–250 Gy, the relative differences in D_E results were on average < 1%, while for samples with $600 < DE < 1000$ Gy, linear conversion yields D_E values on average 1–2% higher than those obtained with OriginPro 9.1[®].

More recently, Joannes-Boyau and Bodin (2014) introduced MCDoseE 1.0, a new fitting program for dose evaluation based on a Monte Carlo computation algorithm. The purpose of the present work is to thoroughly assess the performance of a new version of this program (2.0) to Origin. Several fitting functions were tested on DRCs of fossil tooth enamel and quartz samples (aluminium centre). It is worth mentioning here that although the MCDoseE 2.0 program was first designed for ESR DRCs, it can also potentially be used for luminescence dating purposes, where the same fitting functions are frequently employed. The program is provided in Supplementary Information, with some basic recommendations for its use, to ensure reliable and accurate fitting results.

2. Material and methods

2.1. Samples

To compare both programs, we have used (dose vs ESR intensity) data from a series of tooth enamel and quartz samples that have been processed and analysed according to the standard ESR dating procedures at the Centro Nacional de Investigación sobre la Evolución Humana (CENIEH), Spain (see further details in Duval et al., 2013; Duval et al., 2017).

The DRCs from tooth enamel and quartz samples were obtained using the Multiple Aliquot Additive dose method (Zeller et al., 1967; Duval et al., 2013). Aliquots were irradiated at increasing dose values with either a ⁶⁰Co or a ¹³⁷Cs gamma source. ESR intensities were

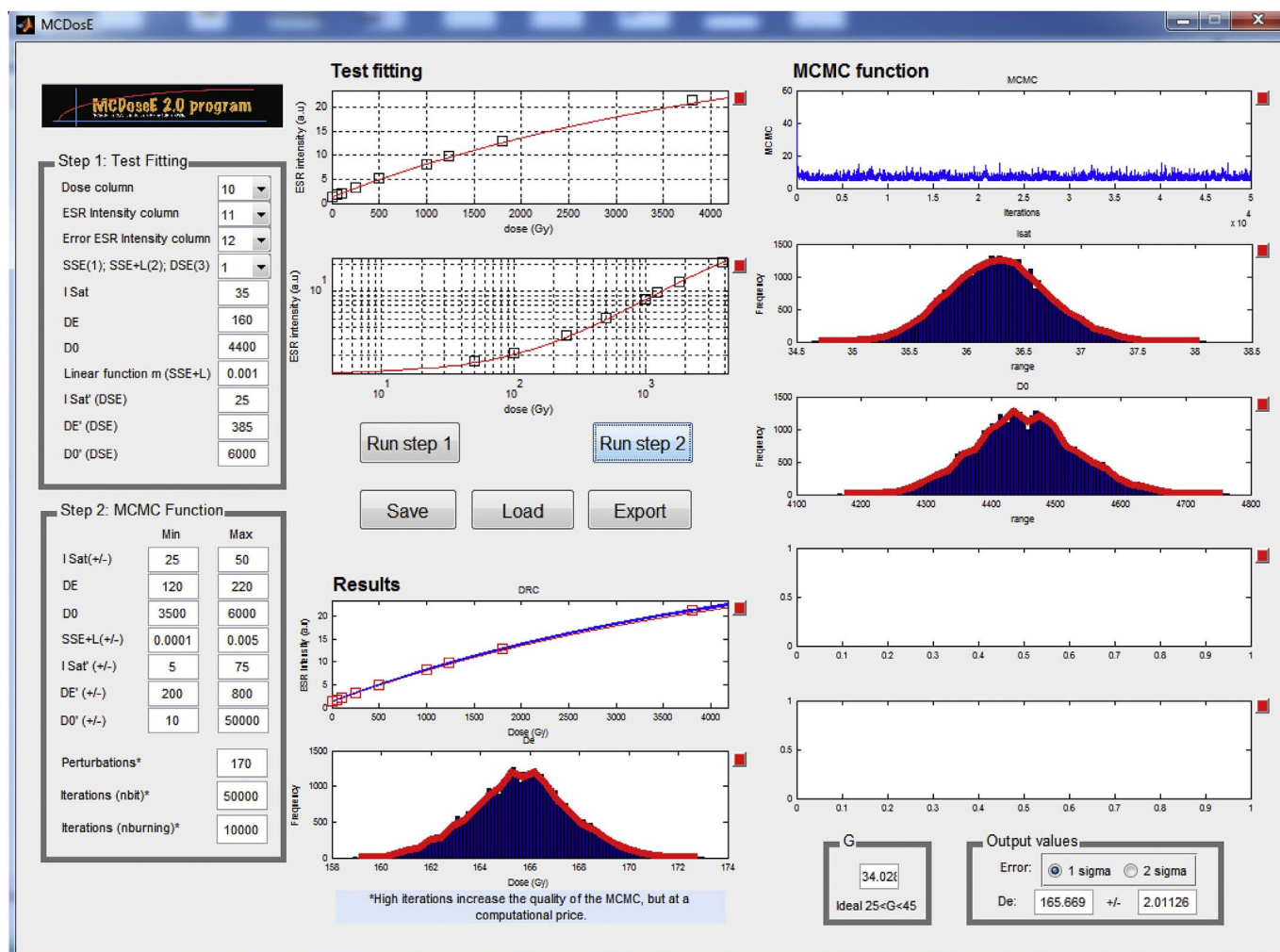


Fig. 1. Screen shot of the MCDoseE 2.0 fitting program (Windows version). Example of SSE fitting results for fossil tooth sample #4 presented in this study.

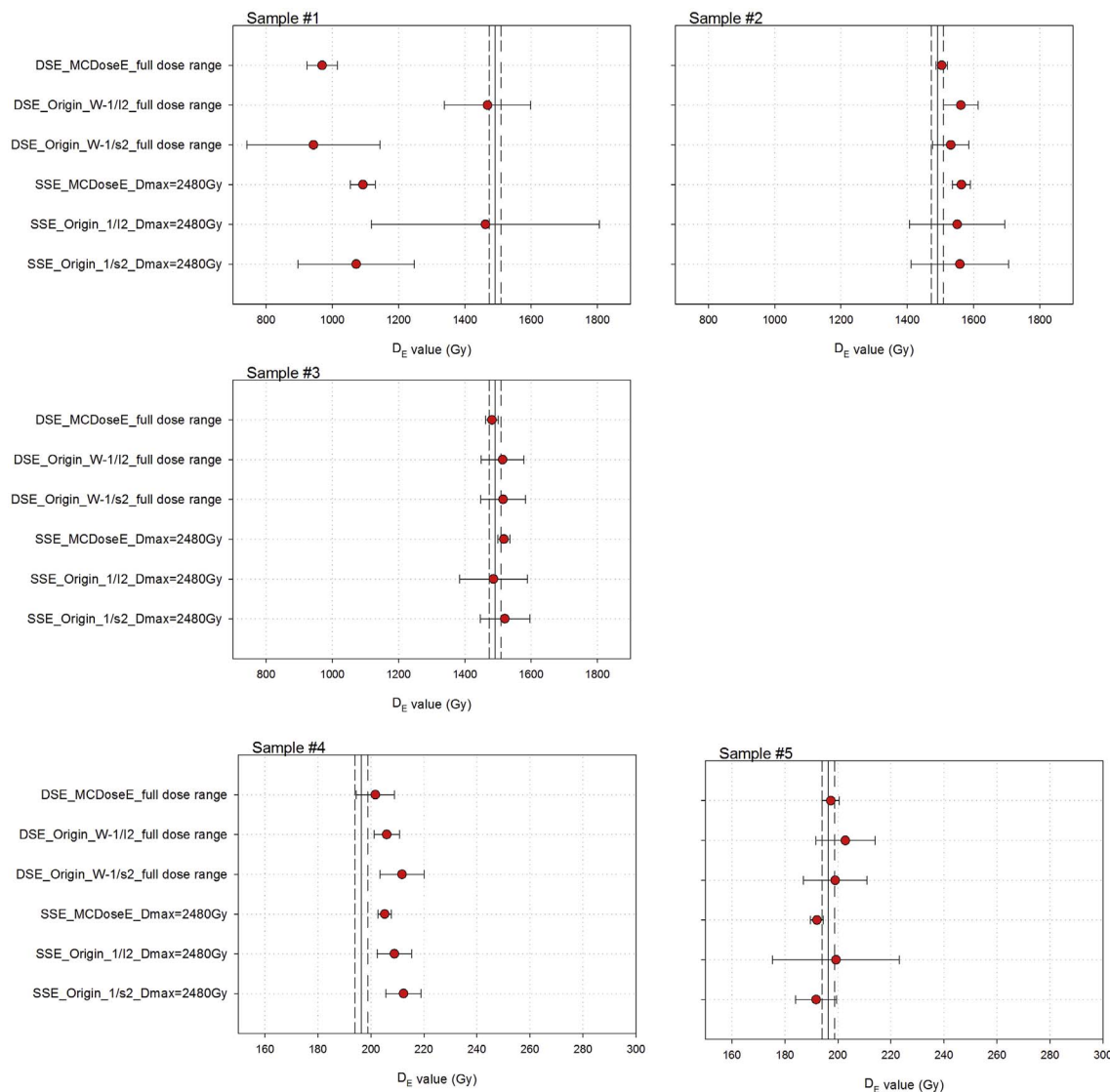


Fig. 2. Comparison of MCDoseE 2.0 and Origin. D_E values obtained by both programmes compared to the known dose given to the enamel samples (sample #1, #2 and #3, at 1491Gy; sample #4 and #5 at 196Gy). The known dose is represented by the thin black line and errors on the dose by the dashed lines. All errors are 1σ.

Table 2

Comparison of the systematic deviations (in %) between the known D_E value (1491Gy for sample #1, #2 and #3; 196Gy for samples #4 and #5) and the recovered D_E using MCDoseE 2.0 and OriginPro 9.1. (data weighted by 1/I² and 1/s²). With the SSE, D_{max} = 2480 Gy for #1, #2 and #3 and 1368 Gy for #4 and #5.

	MCDoseE 2.0		OriginPro			
	DSE	SSE	DSE (1/I ²)	DSE (1/s ²)	SSE (1/I ²)	SSE (1/s ²)
#1	-35.0	-26.8	-1.6	-36.8	-2.0	-28.1
#2	0.9	4.8	4.7	2.7	4.0	4.5
#3	-0.7	1.8	1.5	1.6	-0.4	2.0
#4	2.7	4.5	4.9	7.8	6.4	8.1
#5	0.4	-2.3	3.2	0.2	1.4	-2.4

extracted from the spectra via Bruker WinEPR system software (Version 2.22Rev. 12) with the distance cursor tool. This was done by taking the peak-to-peak amplitude between T1 and B2 of the enamel signal (Grün, 2000) and between the top of the first peak (g = 2.0185) and the bottom of the 16th peak (g = 1.9928) for the Aluminium signal measured in quartz samples (Toyoda and Falgueres, 2003). Each sample was measured at least three times over several days, allowing the calculation of the mean ESR intensity and standard deviation for each

measured aliquot.

The comparison study between the two programs was carried out in two main steps:

- (i) First, we selected five DRCs of enamel samples with known-dose, three with an expected D_E value of 1491 Gy (samples #1, #2, #3 of Duval, 2015) and two with an expected D_E value of 196 Gy (#4, #5; Duval and Grün, 2016).
- (ii) Secondly, we have analysed a set of random samples: 19 tooth samples covering a wide range of D_E between ≈ 100 and ≈ 2500 Gy and 14 quartz samples (Al centre) with D_E values between ≈ 100 and ≈ 2000 Gy.

The tooth enamel DRCs were tested with two fitting functions: a Single Saturating Exponential (SSE) function (with selected D_{max} satisfying the criteria from Duval and Grün (2016)), and a Double Saturating Exponential (DSE) function on the full dose range available. In contrast, the exponential + linear (EXP + LIN) function was tested for the Al centre measured in quartz samples, using the recommendations by Duval (2012). The detailed equations of the fitting functions applied in this work are displayed in Table 1.

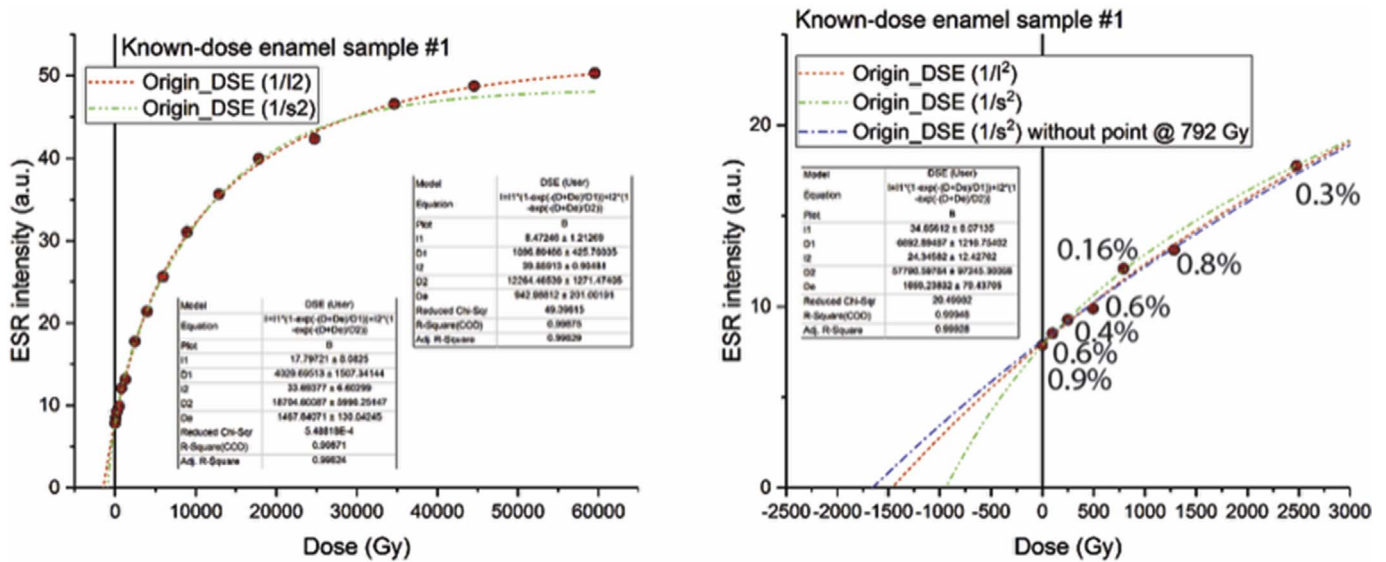


Fig. 3. Impact of measurement precision on the calculated D_E values for the known-dose enamel sample #1. Associated experimental errors are indicated (Right). With OriginPro, the DSE fitting with $1/I^2$ provide the expected D_E within error, whereas the same function with data weighting by $1/s^2$ gives a significantly underestimated result (Left). Because one point has a very small associated error (0.16%) in comparison with the others, it has a huge influence on the fitting results. When this experimental point is removed from the DRC, the D_E derived from the DSE ($1/s^2$) increases by almost 700 Gy, i.e. to 1659 Gy. If this point is given an error of 0.5%, then the D_E remains around 1655 Gy.

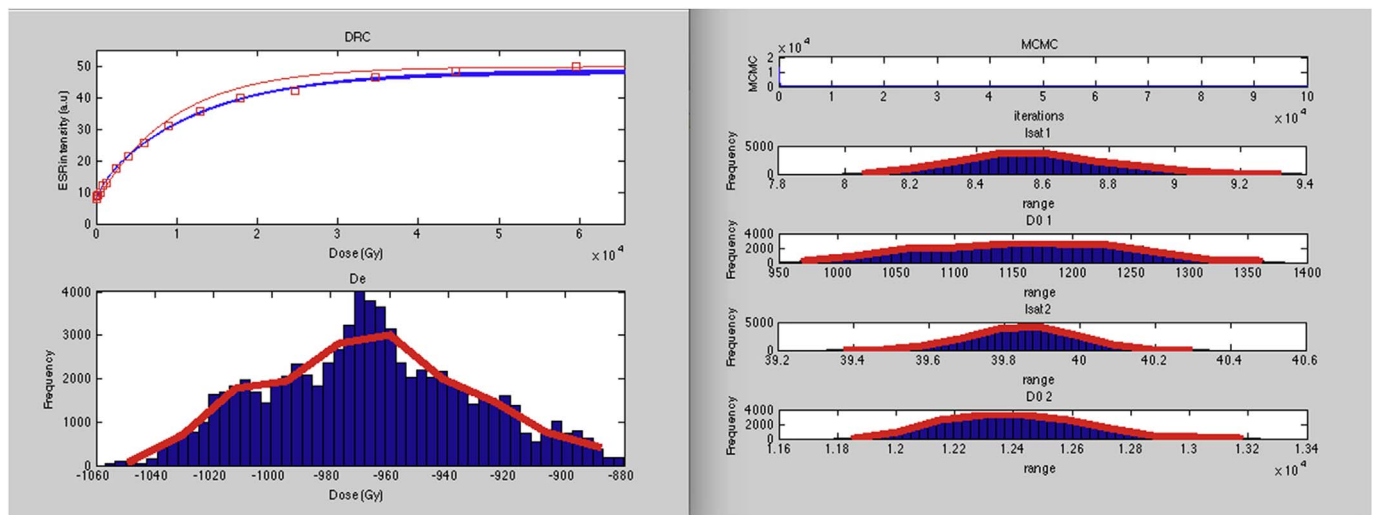


Fig. 4. Results of MCDoseE 2.0 on sample #1. With original experimental errors, iterations of 100,000, on a DSE calculation over the full dataset. We note that even after 100,000 iterations the solution distribution remains chaotic and poorly constrained.

2.2. D_E evaluation with OriginPro software

With OriginPro 9.1, the fitting functions are created as user-defined functions through the Non-linear Curve Fit box. Fitting was performed using two data weighting options: (i) the inverse of the squared ESR intensities ($1/I^2$), which results in giving more weight to the first points of the DRCs (Grün and Brumby, 1994; Duval and Grün, 2016); (ii) the inverse of the squared experimental errors (1-standard deviation) ($1/s^2$), which consists in giving greater importance to the points with the smallest errors, similar to MCDoseE 2.0. With this software, the non-linear fitting is done by an iterative linearized procedure, using a Levenberg-Marquardt (L-M) algorithm by chi-square minimization. OriginPro 9.1 offers also the possibility to use the Simplex algorithm, especially when the initialization of the parameters is complicated. In that case, the Simplex method may be used to get the approximate parameter value for further fitting calculation with the L-M method. Further details about the non-linear fitting performed by OriginPro 9.1 may be found in the Origin 9.1 User Guide (2013).

2.3. MCDoseE 2.0 program description

DRC estimation represents a non-linear regression problem, making it difficult to propagate measurement errors towards the fitted curve uncertainties. The “classical” optimization approaches, e.g. used in OriginPro, suffer from two major shortcomings: (i) the frequent need of data weighting and (ii) the impossibility of accurately propagating the error into the D_E calculation. MCDoseE 2.0 tackles the problem using a non-linear Bayesian framework, where the solution is a probability distribution on the dose equivalent that fully describes the level of knowledge on the solution. The MCDoseE 2.0 program uses a Markov Chain Monte Carlo (MCMC) scheme based on the Metropolis-Hastings algorithm to explore potential solutions (Fig. 1) (Metropolis et al., 1953, see also Joannes-Boyau and Grün, 2011; Joannes-Boyau, 2013 for ESR examples). There is no linearization involved with this algorithm, and this allows for a truthful assessment of uncertainties.

MCDoseE explores the potential solutions using a random walk based on the MCMC algorithm. The estimation starts with a random

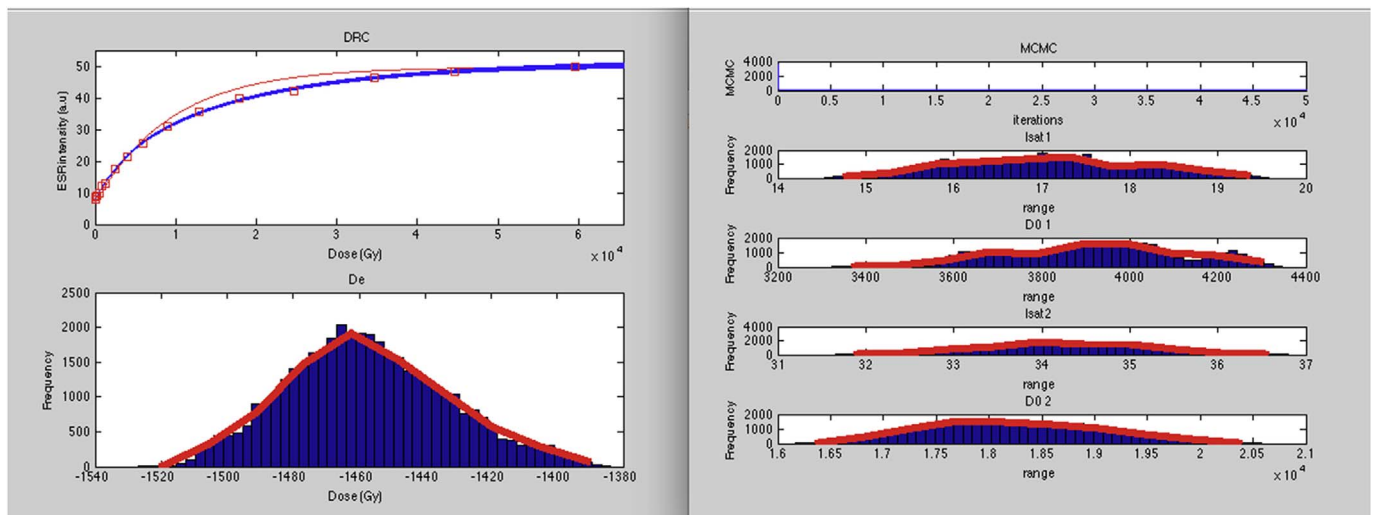


Fig. 5. Results of MCDoseE 2.0 on sample #1. With homogenised associated errors, iterations of only 50,000, on a DSE calculation over the full dataset. If we compare with Fig. 3, we can see that the MCMC is able to accurately calculate a Gaussian distribution when the associated errors are not limiting the solution space.

guess for the DRC within the dimensional space model. Parameters are initialized using reasonable values in order to guide the first fitting. At each step a new DRC is then proposed after perturbing the ESR intensities within their corresponding error. In other words, new values of ESR intensities are randomly generated by the Monte Carlo simulations, within the Gaussian distribution dictated by the mean value and standard deviation at each point. The new DRC is then compared to the previous fitting for goodness of fit and randomly accepted according to the perturbation parameters. This process is iterated until the ensemble of accepted models converges towards a stationary distribution. Detailed instructions can be found in the MCDoseE 2.0 user guide contained in the zip file of the [supplementary material](#).

3. Results and discussion

3.1. Known-dose enamel samples

3.1.1. Comparison between MCDoseE and OriginPro

The D_E values of five known-dose enamel samples were estimated using both OriginPro and MCDoseE 2.0 programs. Results are graphically displayed in Fig. 2 and systematic deviations from the expected dose are given in Table 2. The following observations can be made from samples #2 to #5:

- all the D_E results agree within error, regardless of the program, the fitting function or the data weighting options considered.
- MCDoseE 2.0 recovered the expected dose with the DSE function and in most cases with the SSE. The systematic deviation of MCDoseE 2.0 using the DSE fitting on the full range of the DRC shows the most accurate results and fitting for all solutions, with the errors between the recovered and known D_E to be less than 3% and in most cases 1%.
- The fitting performed with OriginPro and data weighting by $1/s^2$ provides virtually the same results as MCDoseE 2.0. Results derived from data weighting by $1/I^2$ are nevertheless very similar too.
- The precision in the D_E estimates using MCDoseE 2.0 is significantly better than using OriginPro. However, as a consequence, those highly precise results do not systematically overlap with the expected dose, unlike those from OriginPro. This aspect will be further explored in section 3.1.2. It should be mentioned here that the precision of the D_E values provided by OriginPro can be significantly improved by pooling all repeated ESR intensities in a single DRC (e.g., Duval, 2012). This has no impact of the D_E value, but may

reduce the associated error by 50–60% if ESR intensities show very little variation over repeated measurements. This procedure, however, has not been used here, as the comparison between the two programs was intended to be based on exactly the same ESR data inputs.

In contrast, MCDoseE 2.0 using DSE and SSE functions for sample #1 gives an aberrant D_E values of 1092 ± 38 Gy and 969 ± 46 Gy, respectively, which are -27% and -35% lower than the expected value. OriginPro provides exactly the same results using $1/s^2$ data weighting, whereas using $1/I^2$ recovered the expected dose. This indicates there are some leverage effects during the fitting, and some points carry an unexpectedly high weight with $1/s^2$, but not with $1/I^2$. A close look at the ESR data set indicates that the problem originates with only one very small experimental error ($< 0.2\%$), namely the fourth ESR intensity point. The MCDoseE 2.0 program rejects part of the solution through the iteration and burn-in process (burn-in is a colloquial term that describes the practice of throwing away some iterations at the beginning of an MCMC run), to avoid being stuck in a local minimum. A similar issue affects the OriginPro fitting with $1/s^2$, by giving a major weight to this experimental point. As a matter of fact, when this data point is removed from the DRC, the resulting D_E value is of 1659 ± 79 Gy (vs. 943 ± 201 Gy previously), i.e. consistent within 2σ of the expected D_E value (Fig. 3).

To summarize, the first set of results shows that MCDoseE provides results consistent with those derived from OriginPro. For a given function, MCDoseE provides D_E estimates that are closer to the expected value than OriginPro (and is frequently more precise), regardless of data weighting. Yet, the high precision of some of the results by MCDoseE 2.0 prevents overlap of the solution with the expected value. In other words if the uncertainties limit the associated error, the results will somehow be inaccurate. These specific factors driving the D_E precision with MCDoseE 2.0 are explored in the next section.

3.1.2. Testing the impact of measurement precision on the D_E results obtained with MCDoseE

One of the main reasons for creating the MCDoseE program was the ability to meaningfully propagate the experimental error of the ESR intensity measurements. The estimated errors in the Origin software directly depend on the data weighting option selected, and may be biased in a statistical sense. The MCDoseE 2.0 program works differently by considering the input error for each data point within the calculation. By doing so, the program propagates the error to the D_E

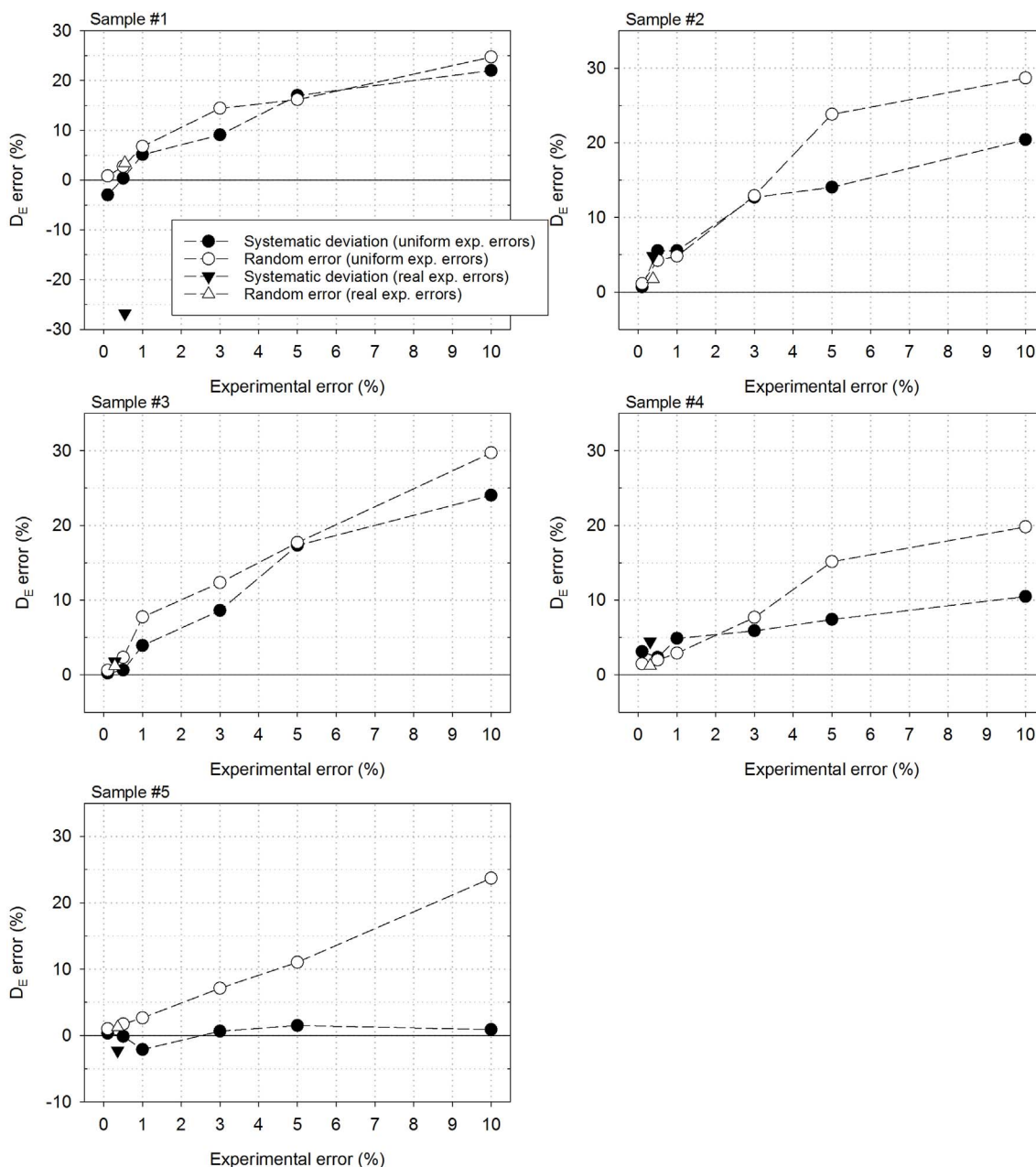


Fig. 6. Impact of the experimental errors on the D_E results (SSE function). Evolution of the systematic error (relative deviation with the expected dose) and random error (D_E precision) with the magnitude of the experimental error of the ESR intensities. Calculations were performed assuming uniform relative experimental errors (circles) for all aliquots of a given sample of 0.1%, 0.5, 1%, 3%, 5% and 10% (1σ errors). The results (1σ errors) for real non-uniform experimental errors (triangle) are also indicated (the mean experimental error derived from the individual error on each aliquot was used for the plot). These tests were run on known-dose samples: for samples #1, #2 and #3, the expected dose is 1491 Gy, while it is 196 Gy for samples #4 and #5.

results, but also influences the solution distribution. In Table 2, sample#1 shows a very large systematic deviation from the known dose. As mentioned above, the reason for such a deviation originates in the associated error of one specific point (at 792Gy, the associated error is $< 0.2\%$). The small error forces all solutions to pass by or not to deviate far from the point, inducing the MCMC to remain stuck on a poor solution corresponding to a local minimum, and ultimately producing a large underestimation of the D_E ($969\text{Gy} \pm 46$) (Fig. 4). By increasing the associated 1σ error of this particular point to 1%, the MCMC is able to fully inspect the distribution, with a D_E solution of $1464\text{Gy} \pm 31$ (Fig. 5). With the increase of the associated error, MCDoseE 2.0 obtains a systematic deviation of -1.8% (see Table 2). It appears therefore extremely important to test the influence of error propagation on the solution.

To do so, we have used the ESR DRCs of 5 known-dose samples, but instead of using the experimental uncertainties, the relative errors were set to specific values. Tests were performed using 0.1%, 0.5%, 1%, 3%, 5% and 10% of 1σ error for the ESR intensities of all aliquots. Figs. 6 and 7 show the influence of the set errors on the overall D_E solution (systematic deviation and random error). As expected, the magnitude of the experimental errors directly impacts the D_E results calculated with MCDoseE 2.0. The higher the error in the ESR intensities, the higher the resultant systematic deviations and random errors (1σ). The results derived from the SSE and DSE functions show the same trends. With experimental errors of 0.1%, the systematic deviation is $< 3\%$, while the precision on the D_E is $< 1\%$. In contrast, with a 10% error, the D_E may be overestimated by up to 30%, while the D_E precision may also reach a similar value.

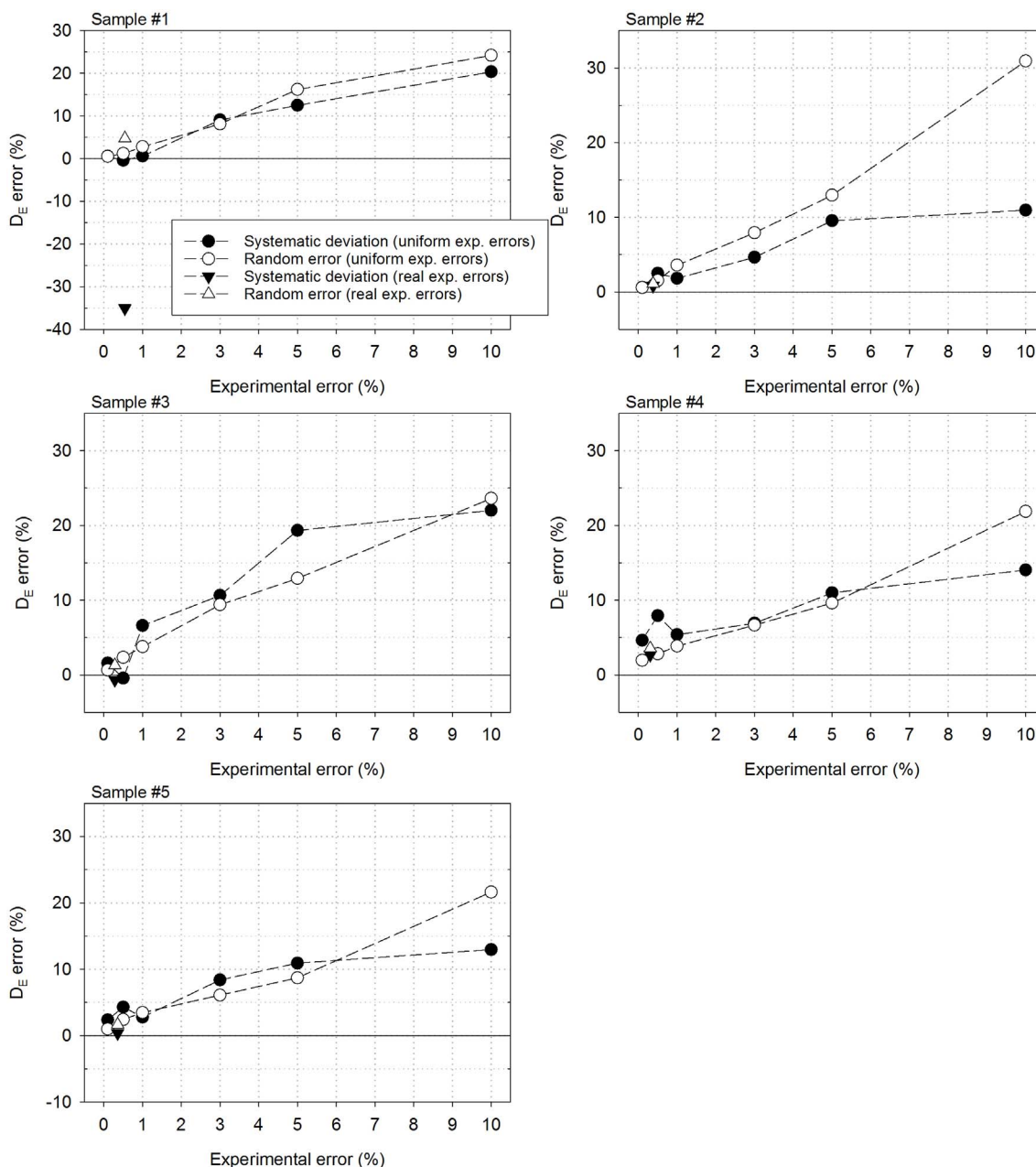


Fig. 7. Impact of the experimental errors on the D_E results (DSE function). Evolution of the systematic error (relative deviation with the expected dose) and random error (D_E precision) with the magnitude of the experimental error of the ESR intensities. Calculations were performed assuming uniform relative experimental errors (circles) for all aliquots of a given sample of 0.1%, 0.5, 1%, 3%, 5% and 10% (1σ errors). The results (1σ errors) for real non-uniform experimental errors (triangles) are also indicated (the mean experimental error derived from the individual error on each aliquot was used for the plot). These tests were run on known-dose samples: for samples #1, #2 and #3, the expected dose is 1491 Gy, while it is 196 Gy for samples #4 and #5.

In general, the 1σ errors on the ESR intensities are around 1% and rarely exceed 3% (Duval et al., 2013). With 1% error on the ESR intensities, MCDoseE 2.0 program provides D_E values that are in agreement within 5% with the expected value, while the precision remains < 10% (1 sigma).

The D_E error derived from the real experimental (and non-uniform) errors is also plotted on Figs. 5 and 6 (triangles) for comparison. Interestingly, those SSE results for 2 of 5 samples follow the trend indicated by the simulations. However, for the remaining 3 samples, the systematic deviation is higher than expected. This is especially striking for sample #1, for which the true D_E is underestimated by > 25% with the SSE (and > 30% with the DSE). This is basically due to the fact that there are one or two points showing a very high precision and which thus drive the fitting of MCDoseE. Given their precision, they carry a

very high weight in the fitting (see Fig. 3). If for any reason, this point significantly deviates from the behaviour shown by the other points, then it will lead to the calculation of an incorrect D_E result. In contrast, when ESR intensities have all similar errors, then fitting provides correct values. For example, with sample #1, one point at 792 Gy shows an extremely precise ESR intensity (0.16%, while the surrounding points have an error > 0.3%: see Fig. 1). When removing the point, MCDoseE 2.0 provides a D_E result of 1572Gy ± 33, i.e. much closer to the true D_E.

In summary, those results indicate that the accuracy of the D_E value calculated by MCDoseE 2.0 is directly dependent on the precision on the ESR intensities (e.g. the higher the precision, the smaller the random error on the D_E). This implies that the correct uncertainties in such points is probably underestimated, possibly due to the influence of

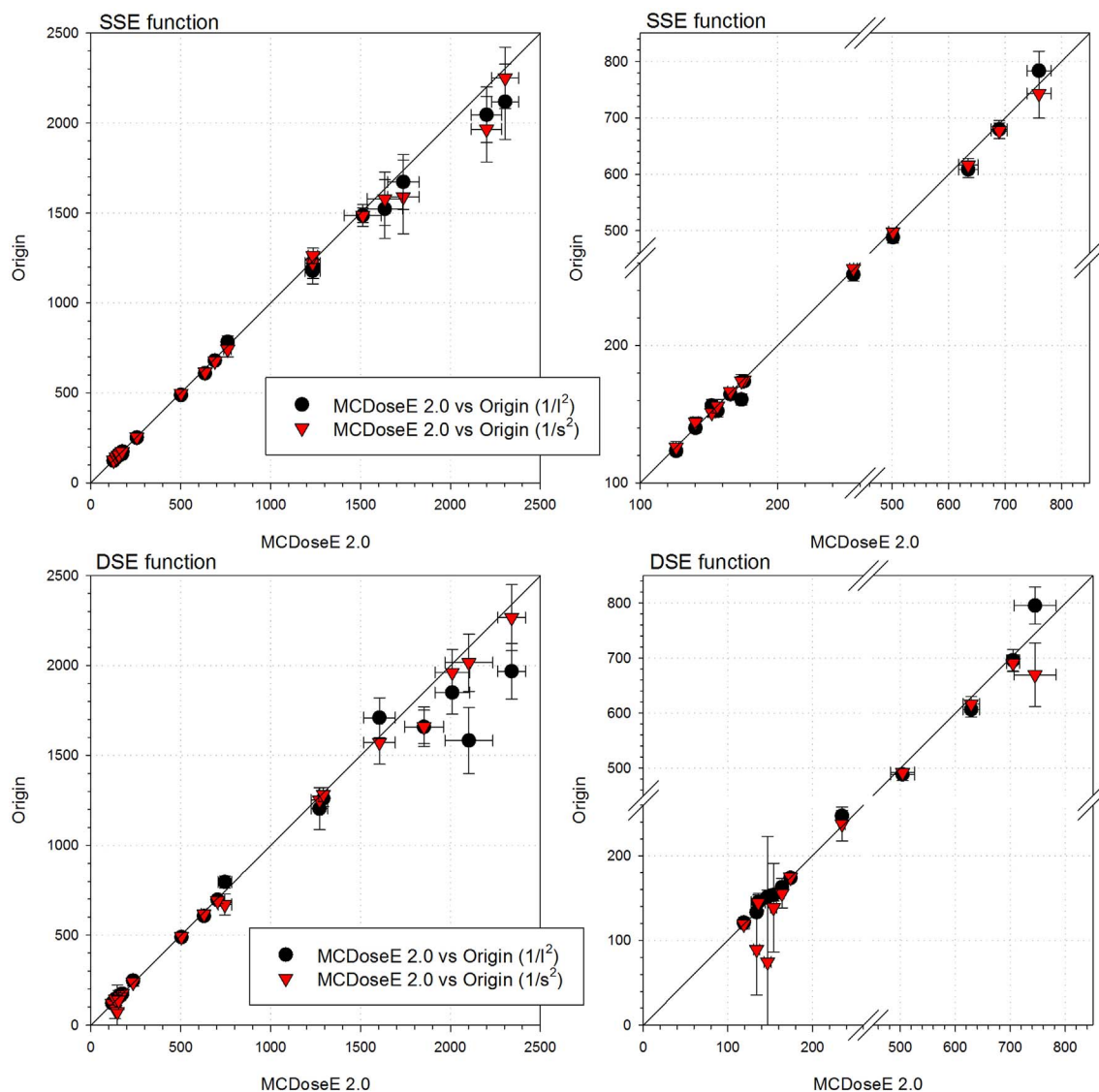


Fig. 8. Comparison of MCDoseE 2.0 and Origin for random fossil enamel samples. Correlation between D_E results obtained with MCDoseE with those from Origin (data weighting by $1/I^2$ and $1/s^2$) for 19 tooth enamel samples, with both DSE and SSE functions fitted. The solid line represents a 1:1 ratio. All errors are 1σ .

systematic error. Such under-reporting of errors is likely to lead to problems, and should be part of the checks. This also provides a potential justification for removing points, or increasing the associated errors. The main limitation of the program remains in the case that one or two points show a significantly higher precision than the group. By default those highly precise value(s) will carry a major weight in the fitting and may potentially induce the calculation of an incorrect D_E . Consequently, we recommend to the users of this program to double check the fitting, by comparing the results obtained using the real experimental errors with the one obtained using uniform errors of 0.5 or 1.0%. Ideally, all D_E values should remain very close to each other. If not, then one or more associated errors might be problematic. MCDoseE 2.0 has a function to quickly change the error to 0.5% of the ESR intensity (for more information see the MCDoseE 2.0 user guide included in the [supplementary file](#) of this paper).

3.2. Random samples

To further compare the two programs, we have tested the calculations on DRCs from random fossil teeth ($n = 19$) and quartz ($n = 14$) samples. It is important to note that the true dose for those samples is unknown, limiting the comparison to the solutions given by the two

programs.

Results obtained for the fossil teeth using the SSE and DSE functions are summarized in [Fig. 8](#). Both programs give statistically indistinguishable D_E results for most samples, with an average ratio of 0.98 and 0.93 for the DSE function with weighting options by $1/I^2$ and $1/s^2$ respectively, and of 0.97 and 0.98 for the SSE with $1/I^2$ and $1/s^2$ respectively. The highest difference observed for the DSE between OriginPro ($1/s^2$) and MCDoseE is caused by the results obtained for two samples (#3 and #5, see [supplementary information table S4](#)): for these samples, OriginPro provides unrealistically small D_E values with large associated errors, suggesting that the fitting results are unreliable. Without those two samples, the D_E ratio between the 2 programs increases to 0.97. In other words, MCDoseE gives on average results that are in agreement within 3% with those derived from OriginPro, regardless of the weighting option considered.

In terms of precision, MCDoseE 2.0 yields average 1σ errors of 2.7% using the SSE, vs 4.2% for Origin (both $1/I^2$ and $1/s^2$). Using the DSE function for 19 samples, the average 1σ errors are 3.5% (MCDoseE) vs 5.0% for Origin ($1/I^2$). In contrast, Origin using $1/s^2$ yielded an unexpectedly large average error of 20.5%. This is the result of the two outliers mentioned above: without these, the mean error drops to 7.7%. In summary, the fitting results indicate that MCDoseE provides a better

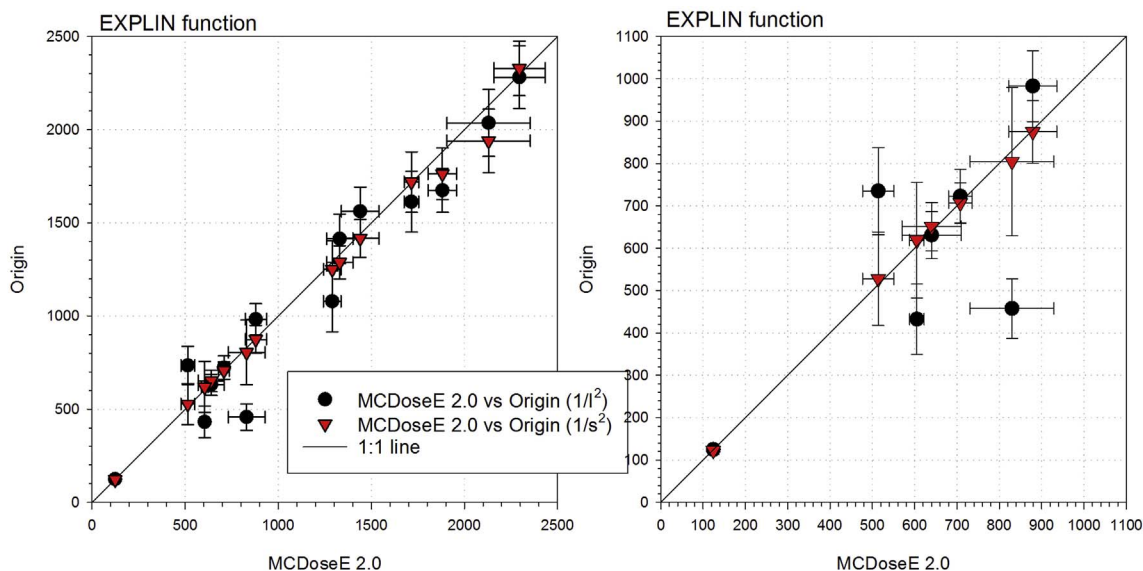


Fig. 9. Comparison of MCDoseE 2.0 and Origin for random quartz samples (Aluminum centre). Correlation between D_E results obtained with MCDoseE with those from Origin (data weighting by $1/I^2$ and $1/s^2$) for 14 quartz samples. Fitting was performed using EXP + LIN function (see Table 1). The solid line represents a 1:1 ratio. All errors are 1σ .

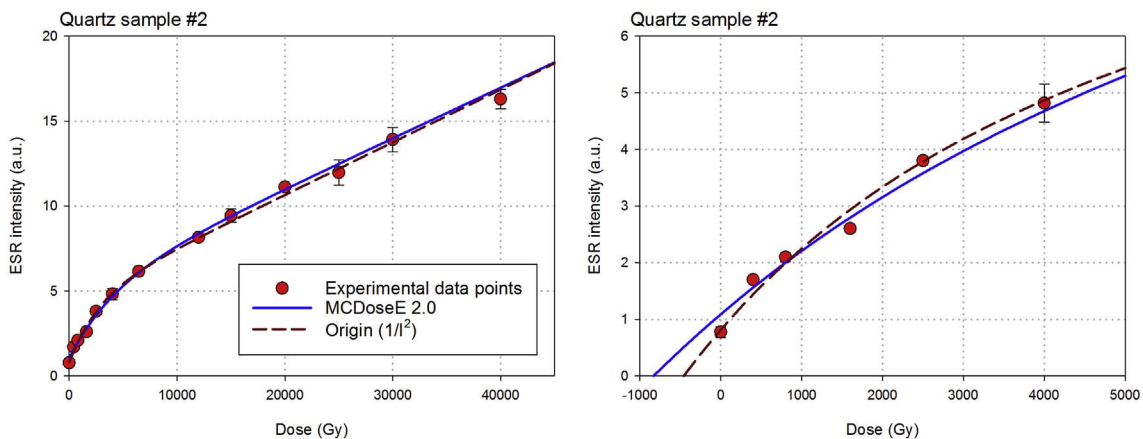


Fig. 10. Comparison of the DRCs obtained by MCDoseE 2.0 and Origin for the quartz sample #2. This graph illustrates the impact of the fitting procedure in case of scattered DRCs.

precision of the D_E value by 30–40% on average compared to Origin. Similar results can be observed with the EXP + LIN function on the quartz samples (Fig. 9). The high correlation between the D_E obtained with each program is visually striking, with an average systematic deviation < 3% for both $1/s^2$ and $1/I^2$. MCDoseE 2.0 and OriginPro ($1/s^2$) provide the closest values, while the $1/I^2$ option provides a very different D_E value (deviation > 20%) for three samples. One example of a sample providing inconsistent results between MCDoseE 2.0 and OriginPro is given in Fig. 10. This discrepancy is simply explained by the very low position of the natural point with respect to the first irradiated value. MCDoseE 2.0 by its algorithm design does not favour the natural point. Even more so in this case as the associated error is relatively high (approx. 12%). Conversely, this same point will carry maximum weight with Origin ($1/I^2$), inducing thus the estimation of a much lower D_E value. It is however impossible to determine here which option provides the right answer, as the true dose of the sample is unknown. That said, it should be mentioned that 90% of the values obtained by both programs on the random samples for both functions are within error. D_E precision using MCDoseE is on average 6.3%, while it is of approx. 11% for OriginPro ($1/I^2$ and $1/s^2$). Consequently, MCDoseE 2.0 provides an increased precision by about 40% in comparison with OriginPro, which is consistent with the previous observation made for the random enamel samples.

4. Conclusion

Both the MCDoseE 2.0 program and Origin software provide highly consistent results. The MCDoseE 2.0 program shows similar capabilities to the Origin fitting. By design, the program puts much emphasis on the experimental errors associated with the ESR intensities. Those errors drive the fitting process and directly impact the correctness and precision of the results. A similar phenomenon occurs when using Origin with data weighting by $1/s^2$. Consequently, the accuracy of the fitting results is also directly limited by the reliability of the experimental errors. The presence of an experimental point with an abnormally small error can potentially induce the MCMC to be stuck in an inaccurate local minimum solution space. To avoid this issue, we recommend that each associated error should be checked to comply with a 2-sigma deviation from the mean error. For a quick evaluation of the potential bias induced by some points showing exceptionally low errors, the user can select a standard 0.5% error in the program for all points, and see whether the result remains within error of the D_E derived from experimental errors. If all parameters are carefully checked and validated, the MCDoseE 2.0 program offers a reliable dose reconstruction fitting procedure for ESR dating protocol.

The results presented here (re)open some old discussions regarding the importance of an appropriate data weighting in the fitting

procedures in ESR dating (e.g. Lyons et al., 1992; Grün and Brumby, 1994). So far, it has mostly been recommended to use a data weighting by $1/I^2$ (Grün and Brumby, 1994; Duval and Grün, 2016), as a leverage to force the fitted curve to go through the natural point. This is by definition based on the consideration that the natural point is the most important in the DRC, and the further in dose are the experimental points, the less weight they should carry. When using data weighting by $1/s^2$ or the MCDoseE program the leverage philosophy is different, as it is then considered that the most precise points in the DRC are the most important, whatever their position. Consequently, depending on the type of data weighting selected, a given experimental point will carry more or less weight and will not influence in the same way the fitting results. This has been demonstrated here with known-dose sample #1, and some random samples (e.g. Figs. 2, 7 and 8). By definition a highly precise point will not impact the fitting by $1/I^2$, contrary to that with $1/s^2$ (Fig. 3). In contrast, an abnormally scattered point in the first dose steps of the DRC will strongly impact the $1/I^2$, but not the $1/s^2$ if its experimental error is similar or higher than those of the other points (Fig. 9).

In any case, it should be mentioned here that if the ESR data set is good (i.e. experimental data points have little scatter, all with similar and small experimental errors), the data weighting option selected should not matter, and all the fitting results should be somewhat similar (see examples in Figs. 2, 8 and 9). As a precaution, and to avoid any potential significant bias induced by an outlier point, we recommend thus the comparison of fitting results derived from each data weighting option. If those results are significantly different, then one or more points may carry an abnormally high weight in the fitting, and a closer look at the ESR data set is then needed to evaluate whether this is justified.

The use of MCDoseE 2.0 will open new avenues worth exploring in the future, such as the meaning of experimental errors associated with the ESR intensity. Some results of this work indicate that a highly precise point is not necessarily correct and may induce a bias in the fitting results. Experimental errors are in most cases derived from the repeated measurement of a given points, but other sources of uncertainty that are most likely involved are usually not considered. For example, it seems clear that the error on the irradiation dose values should at some point also be taken into consideration in the fitting procedure, in order to correctly assess the true error on the D_E . Further work is required in this direction and a new version of MCDoseE will be developed in the future to take this uncertainty into consideration.

Acknowledgement

This research was financially supported by the Australian Research Council (DP140100919), by the Southern Cross University Postdoctoral Research Fellowship and by the Miller Institute for Basic Research at the University of California, Berkeley. Aspects of this research have been funded by an ARC Future Fellowship Grant (FT150100215) awarded to M. Duval. We thank Dr M. Tongue for proof reading and editing the terrible “frenghish” of this manuscript as well as for his valuable comments.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.quageo.2017.11.003>.

References

- Apers, D., Debuyst, R., de Canniere, P., Dejehet, F., Lombard, E., 1981. A criticism of the dating by electronic paramagnetic resonance (ESR) of the stalagmite floors of the Caune de l'Arago at Tautavel. In: *Absolute Dating and Isotope Analyses in Prehistory—methods and Limits*. CNRS, Paris.
- Duval, M., 2012. Dose response curve of the ESR signal of the Aluminum center in quartz grains extracted from sediment. *Anc. TL* 30 (2), 1–9.
- Duval, M., 2015. Evaluating the accuracy of ESR dose determination of pseudo-Early Pleistocene fossil tooth enamel samples using dose recovery tests. *Radiat. Meas.* 79, 24–32.
- Duval, M., Grün, R., 2016. Are published ESR dose assessments on fossil tooth enamel reliable? *Quat. Geochronol.* 31, 19–27.
- Duval, M., Grün, R., Falguères, C., Bahain, J.-J., Dolo, J.-M., 2009. ESR dating of Lower Pleistocene fossil teeth: limits of the single saturating exponential (SSE) function for the equivalent dose determination. *Radiat. Meas.* 44 (5–6), 477–482.
- Duval, M., Guilarte Moreno, V., Grün, R., 2013. ESR dosimetry of fossil enamel: some comments about measurement precision, long-term signal fading and dose–response curve fitting. *Radiat. Prot. Dosim.* 157 (4), 463–476.
- Duval, M., Arnold, L.J., Guilarte, V., Demuro, M., Santonja, M., Pérez-González, A., 2017. Electron spin resonance dating of optically bleached quartz grains from the Middle Palaeolithic site of Cuesta de la Bajada (Spain) using the multiple centres approach. *Quat. Geochronol.* 37, 82–96.
- Grün, R., 2000. Methods of dose determination using ESR spectra of tooth enamel. *Radiat. Meas.* 32 (5–6), 767–772.
- Grün, R., 2002. ESR dose estimation on fossil tooth enamel by fitting the natural spectrum into the irradiated spectra. *Radiat. Meas.* 35 (1), 87–93.
- Grün, R., 2006. A simple method for the rapid assessment of the qualitative ESR response of fossil samples to laboratory irradiation. *Radiat. Meas.* 41 (6), 682–689.
- Grün, R., Brumby, S., 1994. The assessment of errors in past radiation doses extrapolated from ESR/TL dose–response data. *Radiat. Meas.* 23 (2–3), 307–315.
- Han, F., Bahain, J.-J., Boëda, É., Hou, Y., Huang, W., Falguères, C., Rasse, M., Wei, G., Garcia, T., Shao, Q., Yin, G., 2012. Preliminary results of combined ESR/U-series dating of fossil teeth from Longguo cave, China. *Quat. Geochronol.* 10 (0), 436–442.
- Hayes, R., Haskell, E.H., Kenner, G.H., 1998. An assessment of the Levenberg-Marquardt fitting algorithm on saturating exponential data sets. *Anc. TL* 16 (2), 57–62.
- Hoffmann, D., Woda, C., Mangini, A., 2003. Equivalent dose determination in foraminifera: analytical description of the CO₂–signal dose–response curve. *Radiat. Meas.* 37 (1), 95–101.
- Joannes-Boyau, R., 2013. Detailed protocol for accurate non-destructive direct dating of human remains. *Geochronometria* 40, 322–333.
- Joannes-Boyau, R., Bodin, T., 2014. The MCDoseE 1.0 program: a Monte Carlo computation for dose response curve. In: *Proceedings of the 14th International Conference on Luminescence and Electron Spin Resonance Dating*, Montreal (Canada), pp. 92.
- Joannes-Boyau, R., Grün, R., 2011. A comprehensive model for CO₂–radicals in fossil tooth enamel: implications for ESR dating. *Quat. Geochronol.* 6, 82–97.
- Kinoshita, A., Barreto, A., Alves, R., Maria Figueiredo, A., Eduardo de Souza Sarkis, J., Dias, M.L., Baffa, O., 2008. ESR dating of teeth from northeastern Brazilian megafauna. *Radiat. Meas.* 43 (2–6), 809–812.
- Küçükuysal, C., Engin, B., Türkmenoğlu, A.G., Aydaş, C., 2011. ESR dating of calcrete nodules from Bala, Ankara (Turkey): preliminary results. *Appl. Radiat. Isot.* 69 (2), 492–499.
- Lyons, R.G., Brennan, B.J., Hosking, P.L., 1992. Estimation of accumulated dose and its uncertainties: potential pitfalls in curve fitting. *Anc. TL* 10 (3), 42–49.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087–1092.
- Origin 9 User Guide, 2013. OriginLab Corporation. Edition 210 p.
- Toyoda, S., Falguères, C., 2003. The method to represent the ESR signal intensity of the aluminium hole center in quartz for the purpose of dating. *Adv. ESR Appl.* 20, 7–10.
- Zeller, E.J., Levy, P.W., Mattern, P.L., 1967. Geological dating by ESR. In: *Proc. Symp. Radioactive Dating and Method of Low Level Counting*, pp. 531–540.